Research article

Interactions do not only tell us \textit{when}, but can also tell us \textit{how}: Testing process hypotheses by interaction

JOHANN JACOBY* AND KAI SASSENBERG
Knowledge Media Research Center, Tübingen, Federal Republic of Germany

Abstract

Hypotheses about psychological processes are often tested using traditional mediation analysis. This analysis relies on measurement of a transmitting variable. Conducting this analysis has become almost synonymous with examining process hypotheses. An alternative strategy to mediation analysis (the Testing-a-Process-hypothesis-by-an-Interaction Strategy, TPIS) is illustrated here. TPIS is based on a fully experimental design whereby a hypothesized process is tested by an interaction between the hypothesized cause of an effect and a contextual variable. In the interaction term, the contextual variable allows comparison of the causal effect observed when the process is uninterrupted to the effect observed when the process is interrupted. Thus, TPIS translates a theoretical process hypothesis into a statistical interaction hypothesis that uses a fully experimental design to directly examine the hypothesized process. Copyright © 2010 John Wiley & Sons, Ltd.
BARON AND KENNY’S (1986) PRESCRIPTIONS TO TEST A PROCESS HYPOTHESIS

Since Baron and Kenny’s (1986) seminal article on the distinction between moderation and mediation, social psychological research has experienced an invaluable increase in clarity and conceptual precision. The article made it obvious that the practice of using the terms moderation and mediation interchangeably, which was widespread at the time (James, 2008), was a confusing way to describe the role of a third variable regarding an $X \rightarrow Y$ effect. In many cases, it was not clear whether authors included a third variable with the intention to investigate (1) a psychological process – how an independent variable $X$ leads to an effect in a dependent variable $Y$ via a third, transmitting, variable $TV$, or (2) the condition when an $X \rightarrow Y$ effect occurs, and when this $X \rightarrow Y$ effect is reversed in direction or changed in magnitude.

Beyond disambiguating the roles of third variables, Baron and Kenny (1986) also proposed two different study designs and statistical analyses to test process hypotheses (how) and hypotheses about the conditions for an effect (when).

Baron and Kenny’s (1986) approach to process hypotheses, proposed under the label of “mediation analysis,” had actually been discussed years before by Judd and Kenny (1981; see also James & Brett, 1984). However, it was more fully developed, contrasted with moderation analysis, and made accessible to social psychology in Baron and Kenny (1986; see Kenny, 2008). We will refer to Baron and Kenny’s (1986) approach as traditional mediation analysis in the remainder of this article.

To test a process hypothesis, traditional mediation analysis recommends a study design with one manipulated construct (operationalized as the independent variable $X$), one measured dependent construct (operationalized as the dependent variable $Y$), and one measured transmitting construct (operationalized as the transmitting variable $TV$). If the hypothesized process holds, then $X$ should covary with the transmitting variable $TV$ and $TV$ in turn should covary with the dependent variable $Y$. Such a pattern in the empirical variables is consistent with the process hypothesis for the theoretical constructs and therefore empirically supports it (see Sigall & Mills, 1998).

Baron and Kenny (1986) advise the use of three regression models to test mediation. In a first regression model the effect of $X$ on $Y$ (path $c$ in Figure 1), also referred to as the total effect, Baron & Kenny, 1986; Pearl, 2001) is estimated, while $TV$ ($M$ in Baron & Kenny, 1986, and Figure 1) is ignored. Second, regressing $TV$ on $X$ should reveal that there is an effect of $X$ on $TV$ (path $a$ in Figure 1). In the third regression model $Y$ is predicted by $X$ and $TV$ simultaneously. This model should show that $TV$ predicts $Y$ beyond the prediction performed by $X$ (path $b$ in Figure 1). This third regression model should also show that the effect of $X$ on $Y$ (path $c'$ in Figure 1, also referred to as the direct effect, Baron & Kenny, 1986; Pearl, 2001) is smaller when accounting for $TV$.

In essence, traditional mediation analysis is based on counterfactual reasoning that is central to all causal thinking (Sobel, 1996; see Pearl, 2000, 2001). Factually, in the design that Baron and Kenny (1986) prescribe to test a process hypothesis, the process does actually take place. However, statistically controlling for $TV$ simulates the state of affairs if the process had not taken place: a counterfactual state. Comparing $c$ and $c'$ therefore amounts to comparing the factual, observed state and a counterfactual, simulated state. Both states are considered based on all observations. If the effect of $X$ on $Y$ is substantially larger in the factual state than in the counterfactual state, then the process hypothesis for the $X \rightarrow Y$ effect appears viable.

To statistically test the difference between $c$ and $c'$, Baron and Kenny (1986) recommend the Sobel (1982) test. This test indirectly assesses the difference between paths $a$ and $b$, its conceptual equivalent (for more recent approaches, see MacKinnon, Warsi, & Dwyer, 1995; MacKinnon, Krull, & Lockwood, 2000; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Muller, Judd, & Yzerbyt, 2005; Muller, Yzerbyt, & Judd, 2008; Preacher & Hayes, 2008a; Shrout & Bolger, 2002). Thus, according to Baron and Kenny (1986), the mediation approach for testing a process hypothesis is defined by an experimental design with a measured transmitting variable, a comparison of factual and counterfactual states, and a particular pattern of results yielded by three regression models and test of the product of two regression coefficients (see Table 1).

Moderation is defined by Baron and Kenny (1986) as a hypothesis regarding conditions under which an independent construct influences a dependent construct. Thus, the influence of an independent construct on a dependent construct is thought to depend on the state of a second independent construct represented by a variable $D$. The study design for moderation thus requires two independent variables and one dependent variable. The statistical test of whether the effect of $X$ on $Y$ differs across levels of $D$ is the assessment of a statistical interaction between the two independent variables (Baron & Kenny, 1986). Unlike in mediation analysis, there is no specific causal logic inherent to this combination of hypothesis, operationalization, analysis, and pattern of results.

Thus, Baron and Kenny (1986) suggest specific combinations of causal logic, study design, and statistical analysis to test hypotheses of mediation and moderation. TPIS does not contradict these recommendations or suggest their modification. Rather, TPIS uses a different combination of a causal logic, study design, and statistical analysis to offer an additional strategy for testing a process hypothesis (see Table 1). Thus, for the sake of clarity, we will use the terms mediation and moderation and their derivatives only when we refer to the specific combinations defined by Baron and Kenny (1986). Otherwise we will use the more general terms we have already employed (e.g., transmitting variable).

Why should one care about an additional strategy to test process hypotheses, given that traditional mediation analysis is well understood and widely used? After all, mediation analysis is practical and has become a standard in the literature.
Table 1. Mediation (according to Baron & Kenny, 1986) and TPIS as combinations of psychological hypotheses, underlying causal logic, study design, and statistical analysis strategy, both testing process hypotheses (see text for elaboration)

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Mediation (Baron &amp; Kenny, 1986)</th>
<th>TPIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychological hypothesis</td>
<td>A dependent construct affects an independent construct via a particular process</td>
<td>Comparison of two factual states</td>
</tr>
<tr>
<td>Causal logic</td>
<td>Comparison of factual and counterfactual states of affairs</td>
<td>Intervention: letting a process run uninterrupted versus interrupting it</td>
</tr>
<tr>
<td>Observation of a process-relevant variable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td>IV $X$, manipulated between ps</td>
<td>IV $X$, manipulated between ps</td>
</tr>
<tr>
<td>IV $T_V$, measured</td>
<td>IV $B$, manipulated between ps</td>
<td>IV $Y$, measured</td>
</tr>
<tr>
<td>DV $Y$, measured</td>
<td>DV $Y$, measured</td>
<td></td>
</tr>
<tr>
<td>Statistical analysis</td>
<td>Three regression models</td>
<td>One regression model</td>
</tr>
<tr>
<td>$Y = i_0 + cX + e_0$</td>
<td>$Y = i + qX + rB + sXB + e$</td>
<td></td>
</tr>
<tr>
<td>$TV = i_1 + aX + e_1$</td>
<td>Test of $s$ against 0</td>
<td></td>
</tr>
<tr>
<td>$Y = i_2 + c'X + bTV + e_2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test of $(c - c') \approx (a \times b)$ against 0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: IV, independent variable; DV, dependent variable; TV, transmitting variable; ps, participants.

(Kashy, Donnellan, Ackerman, & Russell, 2009). It also brought an invaluable increase in clarity regarding the terms mediation and moderation and even regarding the nuts and bolts of data collection and analysis (James, 2008, see also Holmbeck, 1997). One simple answer is that having an additional strategy for testing process hypotheses is better than not having an additional strategy. For example, it may be very difficult or even impossible (with current measurement methods) to perform measurement of a hypothesized $TV$ (e.g., cognitive effort spent on a deliberation, see also Spencer et al., 2005) while interfering with that $TV$ may be easier to accomplish. Then it is naturally useful to have a strategy available that does not rely on such measurement. TPIS is such a strategy.

However, there is also a more compelling answer. In some cases, the use of TPIS circumvents difficulties in traditional mediation analysis. We focus on this reason for supplementing traditional mediation analysis with TPIS.

PITFALLS IN TRADITIONAL MEDIATION ANALYSIS

Although traditional mediation analysis has many advantages, its difficulties have repeatedly been discussed (e.g., Hayes, 2009; Preacher & Hayes, 2008b; Sigall & Mills, 1998; Spencer et al., 2005; Zanna, 2004; see also Kraemer et al., 2008). Here, we focus on two particular difficulties that, if acute, might make TPIS an attractive alternative to traditional mediation analysis.

Causal Relation between the Mediator and the Dependent Variable

The fact that the mediator variable is measured and not manipulated in traditional mediation analysis complicates the inference regarding the causal direction of an association between the mediator and $Y$. Consider a simple correlation between two measured variables. The notion that this correlation could establish a causal effect between these variables is dubious. In traditional mediation analysis, statistically controlling for $X$ while estimating the prediction of $Y$ by $TV$ (path $b$) attenuates this problem somewhat, but still leaves a researcher to infer causality from correlation (see also Kenny, Kashy, & Bolger, 1998; Sobel, 2008; Spencer et al., 2005). In this vein, Stone-Romero and Rosopa (2008) have gone as far as claiming that studies in which the proposed mediator is measured are thwarted by their “inability […] to establish the causal sequence” (p. 347) between $TV$ and $Y$. We can only speculate why traditional mediation analysis seems to be overused (Spencer et al., 2005) without regard for this problem in some cases. Perhaps the originally prescribed temporal sequence of measurement after manipulation of $X$ (i.e., first $TV$, then $Y$; Kenny et al., 1998) deceptively suggests that this ordering of measurements necessarily establishes causal direction between psychological events. This may lead to an erroneous dismissal of the problem of causal direction between $M$ and $Y$ (Mathieu & Taylor, 2006).

Interference of the Measurement Itself

The act of measuring the transmitting variable itself may interfere with the very process one investigates (e.g., Sigall & Mills, 1998). Put in a more general way, the process may be altered as soon as it is observed. This fundamental problem of any empirical science is especially virulent if the object of investigation is not directly observable, but must be inferred using sophisticated measurement methods. The very act of observation can change the object of observation. In experimental social psychology it is widely recognized that measurement administration can act as a manipulation (e.g., Schwarz, Strack, Hippler, & Bishop, 1991). This insight leads researchers to carefully select, plan, and even systematically vary the sequence of different measurements within the same study. They do so in order to avoid and control, for example, priming effects (Bargh, 1984; Bargh & Chartrand, 2000), the salience of issues that would otherwise not have been salient were it not for the fact that these issues were artificially brought to attention by items in a preceding self-report or other measure (Sigall & Michela, 1976), or interpretation and context effects (see Schwarz, 1995; Sperber & Wilson, 1995).

More specifically, the act of measuring a potential mediator may alter a process in three ways: First, the measurement may...
interrupt the process, leading to the elimination of the original effect of $X$ on $Y$. For example, the expression of anger during its measurement as a potentially transmitting variable could attenuate the effect of anger on a dependent variable such as punishment. While a systematic difference in the transmitting variable may be found using a self-report measure, the difference in the corresponding transmitting construct may be eliminated by the measurement of $TV$. Therefore the effect on the dependent construct is diluted. It would, however, have been found had the transmitting variable not been measured (see Spencer et al., 2005).

Second, the measurement of $TV$ may be a prerequisite for the effect of $X$ on $Y$ to occur in the first place. In other words, this effect may only emerge if the measurement of $TV$ is administered, but not otherwise. This might happen because the measurement of the variable $TV$ increases participants' awareness of the (manipulated) state of $X$. This awareness might be necessary for the effect of $X$ on $Y$ to occur. Without this awareness the effect of $X$ on $Y$ would not emerge. For example, under certain circumstances, ease of retrieval effects are eliminated once the traditionally associated manipulation check is not administered (Kühnen, 2010).

Third, the measurement of a $TV$ may actually alter the process one aims to examine. For example, the mere act of retrieving information about the self-concept to answer a self-report measure hypothesized to tap a mediating construct may induce a particular mindset or make information accessible that would not come to mind as a simple result of the independent variable. Thus, the measurement of a $TV$ may redirect the effect of the independent variable or introduce a competing effect. Both of these possibilities may alter the process by which the independent variable affects the dependent variable.

**TESTING THE PROCESS WITHOUT MEASURING THE TRANSMITTING VARIABLE**

Some approaches have been proposed to test process hypotheses in a way that alleviates the causal ambiguity in a correlation between $TV$ and $Y$. One is what Spencer et al. (2005) refer to as an *Experimental-causal-chain design* (see also Stone-Romero & Rosopa, 2008). This strategy requires the completion of two different experiments. In Experiment 1, $X$ is manipulated and $TV$ is measured; in Experiment 2, $TV$ is manipulated and $Y$ is measured. If both experiments yield substantial effects, the process path (i.e., the process) from $X$ to $Y$ via $TV$ is established. This Experimental-causal-chain design requires that the measurement (in Experiment 1) and manipulation (in Experiment 2) of the same transmitting construct is feasible. Therefore, while facilitating causal inference, this strategy is not always possible or advisable. As outlined above, there are instances when the transmitting construct cannot be measured or there are good reasons to not measure it. In addition, the Experimental-causal-chain design presumes that one is able to manipulate $TV$ in Experiment 2 to levels comparable to those produced by $X$ in Experiment 1 (see Kenny, 2008; Word, Zanna, & Cooper, 1974). This is difficult to achieve and difficult to confirm empirically. Of course, another drawback to the Experimental-causal-chain design is that the two required experiments take twice the resources as traditional mediation analysis. Given the drawbacks of traditional mediation analysis and the Experimental-causal-chain design alternative, we think that it is worth considering TPIS as another option.

Baron and Kenny (1986) have foreshadowed TPIS in mentioning instances in which “a moderator variable has been useful in suggesting a possible mediator variable” (p. 1178). Instead of speculating about processes after a study, a design and analysis targeting an interaction effect can be considered *a priori* as a test of a process hypothesis. This is the core idea of TPIS. We will illustrate TPIS using published studies. Many studies that manipulate cognitive load (e.g., see DeSteno, Bartlett, Braverman, & Salovey, 2002; Galinsky & Moskowitz, 2007; Gilbert & Krull, 1988; Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008; Schneider & Shiffrin, 1977; Swann, Hixon, Stein-Seroussi, & Gilbert, 1990) and misattribution (e.g., Cooper, 1998; Fried & Aronson, 1995; Hassebrauck, 1987; Cooper, Zanna, & Taves, 1978) use logic similar to TPIS. Other manipulations that follow the logic of TPIS include self-affirmation (e.g., Aronson, Cohen, & Nair, 1999; Sherman & Cohen, 2006), control (Warburton, Williams, & Cairns, 2006), need for uniqueness (Imhoff & Erb, 2009), anticipated anxiety (Greenberg, Martens, Jonas, Eisenstadt, Pyszczynski, & Solomon, 2003), motor simulation (Topolinski & Strack, 2009), and norm salience (Jonas, Martens, Niesta Kayser, Fritsche, Sullivan, & Greenberg, 2008). After the two illustrative examples we present TPIS in detail and make it obvious that it is epistemologically equivalent with testing a process hypothesis with traditional mediation analysis.

**EXAMPLES OF TPIS**

Valdesolo and DeSteno (2008)

Valdesolo and DeSteno (2008) recently examined the moral hypocrisy effect: participants judge a transgression committed by someone else as worse than if it is committed by themselves. Valdesolo and DeSteno discuss whether this effect is the result of spontaneous automatic bias or a more deliberate process in which one’s own transgression is rationalized and justified. In line with the latter possibility, their theoretical account posits the following process: the transgressor (X; target: other vs. self) affects the extent to which a moral judgment is elaborated on in a controlled fashion (TV: elaboration intensity), and this elaboration determines morality judgments (Y). While traditional mediation analysis comes to mind immediately to test this account, Valdesolo and DeSteno (2008) test this process account in a $2 \times 2$ factorial design with two manipulated variables that are examined statistically via an interaction term. This study, while representing a prime example of TPIS, does not make it explicit that this interaction effect is evidence for a process hypothesis equivalent with evidence from a traditional mediation analysis.

According to their process hypothesis, an other’s transgression is followed by less effortful elaboration (TV) than one’s
own. The effect of this difference in elaboration intensity is a difference in rated morality. But for the difference in elaboration to occur, cognitive resources must be available. Thus, a manipulation of cognitive load (variable B), forcing participants to concurrently attend to a digit-string memory task, should eliminate the elaboration difference between the conditions of X (other vs. self). As a result, the difference in morality judgments found under normal circumstances should be eliminated under cognitive load. This is exactly what Valdesolo and DeSteno (2008) found: If participants were not under cognitive load, the hypocrisy effect was found. But under cognitive load there was virtually no difference between the conditions of X (other vs. self). Together, these results form a statistical interaction pattern between X and the cognitive load manipulation B. Thus, the relationship between the focal independent variable (target), and the dependent variable (judged morality) was eliminated because circumstances were created (cognitive load) that did not allow for the process hypothesized to be responsible for that relationship to occur.2 Clearly, this study tells us something substantial about the process underlying the moral hypocrisy effect, even though the transmitting variable has not been measured and no traditional mediation analysis has been conducted. Rather, the researchers manipulated whether the hypothesized process could occur uninterrupted (in the no load condition) or it was interrupted (in the load condition).

Zanna and Cooper (1974)

Zanna and Cooper (1974) tested a central process account of dissonance effects (Festinger & Carlsmith, 1959). According to the original cognitive dissonance account, a discrepancy between one’s attitudes and (seemingly) freely chosen behavior causes an unpleasant arousal. Participants seek to attenuate that arousal by changing their attitude such that it better fits with their behavior. If participants perceive that they had no choice in their behavior, however, the behavior can be easily attributed to external causes, and thus does not contradict attitudes. According to this hypothesis, participants under low choice therefore experience lower arousal after attitude-inconsistent behavior than those under free choice. Thus, the process hypothesis is that perceived freedom of choice (X) has an effect on arousal (TV), and arousal has an effect on attitude change (Y). In other words, the effect of perceived freedom of choice on attitude change is transmitted by arousal.

In their study, Zanna and Cooper (1974) did not measure arousal (the crucial TV) as would have been required in traditional mediation analysis. Rather the authors experimentally manipulated whether participants could attribute their felt arousal to external circumstances, thus rendering the arousal irrelevant as an indicator of attitude-behavior discrepancy.

They did so by either (a) giving participants a placebo pill and telling them that this pill would cause arousal or (b) not giving them a pill (i.e., manipulation B: pill vs. no pill).3 In the no pill condition, the freedom of choice factor should have an effect on arousal, which in turn should result in attitude change. Participants in the arousing pill condition with perceived freedom of choice should also feel the arousal. However, since this arousal can be attributed to the pill rather than to the discrepancy between attitude and behavior, it should not have any bearing on attitude change. The amount of arousal relevant for attitude-behavior consistency in the free choice/pill condition on the one hand and the no free choice/pill condition on the other should thus be equal and therefore cannot predict attitude change. Indeed, this is what the authors found: an effect of freedom of choice on attitude in the no pill condition, but virtually no effect in the arousing pill condition. In essence: the path from X to TV (arousal) was intact in all conditions, but the path from TV to Y and thereby the entire effect of X on Y via TV was uninterrupted versus interrupted by the experimental pill manipulation.

REFORMULATION OF A PROCESS HYPOTHESIS INTO A PREDICTION OF STATISTICAL INTERACTION

The preceding examples demonstrate that the TPIS approach essentially capitalizes on the fact that the causal role of a transmitting variable can be shown by intervention (rather than observation as in traditional mediation analysis; Pearl, 2000). This intervention consists in interrupting the causal path between X and Y via TV. Below, we demonstrate the apparent similarity in the epistemological content of traditional mediation analysis (based on observation) and the TPIS approach (based on intervention). This will also allow researchers to decide whether a process hypothesis that they intend to test, about any particular constructs and variables, can be tested using TPIS (i.e., in a full experimental design and via a statistical interaction term).

The traditional approach to test mediation distinguishes two states of affairs during the analysis:4

(I) The effect of X on Y, c, is computed while ignoring TV. This effect c is a factual, empirically observed difference between means in Y as a function of the difference between the conditions X1 and X2 (which make up the variable X) when TV is left free to vary:

\[ c = \bar{Y}_{X_2} - \bar{Y}_{X_1} \]

where \( \bar{Y}_{X_i} = \text{mean of } Y \text{ in condition } i \) of X.5

2Note that cognitive load (B) is not the transmitting variable (TV) here. Cognitive load is a manipulation of whether X can freely influence TV. TV on the other hand is elaboration intensity. While the cognitive load manipulation consists of adding elaboration demands, elaboration intensity refers to the effort actually spent on particular cognitive operations. These two constructs are related and their corresponding variables may be ecologically correlated, but they are not conceptually identical. Thus, Valdesolo and DeSteno do not simply treat a traditional mediator (which, according to the traditional strategy, should be measured) as a moderator.

3Zanna and Cooper’s (1974) design contained a third condition in which the pill was said to cause relaxation. Although this condition is not discussed here, the theoretical idea behind this condition, the corresponding predictions, and the empirical results are consistent with the causal logic underlying TPIS.

4These two situations correspond to the two effects c and c’ resulting from the estimation of the parameters in two of the three regression models prescribed by Baron and Kenny (1986). They figure prominently at specific points in the four steps prescribed by Baron and Kenny (1986). The difference between c and c’ (i.e., the indirect effect), is the exhaustive core of traditional mediation analysis and fully redundant in meaning with paths a and b in the form of their product (Hayes, 2009; Kenny et al., 1998; Sobel, 2008). Therefore, everything that can be known about mediation in its traditional form is parsimoniously summarized in the difference (c – c’).

Copyright © 2010 John Wiley & Sons, Ltd.

(II) TV is entered as a predictor of Y simultaneously with X, acting as a covariate. In other words, the effect of X on Y is statistically adjusted for the influence of TV on Y (i.e., computed controlling for TV). When the covariate TV is included in the model, the resulting effect of X on Y is \( c' \), which simulates the influence of X on Y if TV had not influenced Y, and thus the process under investigation had been interrupted – a counterfactual state of affairs. This remaining effect of X on Y is

\[
c' = \overline{Y}_{X_2,TV} - \overline{Y}_{X_1,TV}
\]

where \( \overline{Y}_{X_i,TV} \) = mean of Y in condition i of X while statistically controlling for TV across the entire sample.

The term control is used here in a statistical, not an experimental sense. This means that, while the association between TV and Y is mathematically partitioned out in (II), the effect of TV on Y still did in fact occur during the data collection for the study (an important point in the logic of controlling covariates which is not entirely unproblematic, see Miller & Chapman, 2001). In this sense, situation (II) is counterfactual.

There is, however, another meaning of control, germane to the logic of experimental design: control by manipulation. In TPIS, such experimental control is achieved by distinguishing two situations (I′) and (II′) in a fully experimental setting. Situations (I′) and (II′) capture circumstances equivalent to those in (I) and (II) above with the important difference that they are both factual. Specifically, control is exerted by randomly assigning observation units to either of two conditions of a between participants design. These conditions correspond to (I′) and (II′). In the condition corresponding to (I′) the hypothesized process is left uninterrupted. In (II′), in contrast, the hypothesized process will be interrupted by an actual intervention rather than removed statistically.

(I′) The process from X to Y is uninterrupted: X is manipulated and Y is measured, no further measurement or manipulation takes place. This is conceptually equivalent to situation (I) above in which the effect c is assessed without considering TV. In this condition, the effect of X on Y represents the total effect when the process under investigation is left uninterrupted: TV may be influenced by X and, in turn, can influence Y:

\[
c^* = \overline{Y}_{X_2} - \overline{Y}_{X_1}
\]

In situation (I′) the effect of X on Y can be observed as it actually and “naturally” occurs, without the potential for interference from measurement activities. Thus, the concern regarding measurement interference in traditional mediation analysis discussed above is mitigated. Alteration of the process itself by measurement activity of a potentially transmitting variable is entirely impossible here. This condition is otherwise very similar to situation (I) in traditional mediation analysis.

(II′) In this situation, Y is assessed after manipulation of X, but the hypothesized process is interrupted, within a different set of observation units than in (I′). An experimental manipulation is applied which eliminates part or all of the systematic variance from X that is transmitted through TV to Y. The hypothesized causal path between X and Y via TV is now interrupted and the mediating process cannot occur. The remaining effect of X on Y is:

\[
c^{**} = \overline{Y}_{X_2,\text{interrupted}} - \overline{Y}_{X_1,\text{interrupted}}
\]

where \( \overline{Y}_{X_i,\text{interrupted}} \) = mean of Y in condition i of X when the hypothesized process has been experimentally interrupted.

In this situation (II′), the interrupted process condition, we do not rely on statistical control for a measured TV, but only measure Y when, in fact, the hypothesized process cannot have taken place. In traditional mediation analysis, the variance in the observed values of TV is statistically removed, and thereby its systematic transmission into Y is simulated to be absent. In TPIS, however, the transmission of systematic variance in TV is prevented by manipulation. In other words, TPIS does not mathematically simulate the counterfactual state of affairs if the hypothesized process had not taken place (as is the case in the traditional approach employing covariate analysis), but it directly assesses this state of affairs where the hypothesized process in fact did not take place (or took place only to a reduced extent). It thereby aims to alter the process in question. However, other than in traditional mediation analysis where this alteration could be effected by measurement activities, this alteration here is not extraneous to the process at hand, but directly targets the focused process. Since this alteration by intervention is present in one condition (II′), but not in the other (I′), it can be explicitly and directly modeled and tested. A comparison of expressions (2) and (4) confirms that from a causal perspective (II) and (II′) hold the same conceptual content.

Whether the causal path via TV is interrupted, as in (II′), or not, as in (I′), can be encoded in a new variable, B. One may set B = −1 if the process is uninterrupted: situation(I′), and B = +1 if the hypothesized process is interrupted by manipulation: situation (II′).

Then testing the process account means testing the interaction of X and B on Y and this is the reason why this approach is called the Testing-a-Process-hypothesis-by-an-Interaction Strategy. In other words, if the process account is viable, the effect of X on Y will be different conditionally on B. Under B = −1, a positive effect c′ of X on Y should be observed, whereas under B = 1 the effect c^{**} should be nil or substantially smaller than c′ (see Figure 2). The statistical test for the difference between the two effects c′ and c^{**} is the test of the interaction term. It is equivalent to the famous Sobel (1982) test, which tests the difference between c and c′ indirectly via the product of a and b in the traditional mediation test approach.7

7In the present discussion we presuppose a manipulated dichotomous variable X. All involved variables are assumed to be coded such that c has a positive sign. Our argument throughout this paper can be generalized mutatis mutandis to negative effects.

7In traditional mediation analysis, the two situations are rather different analyses on the same set of observations, corresponding to a within participants design.

8Of course the mere omnibus test for the interaction effect is not enough to fully test the predicted pattern (c′ > c′) in a focused fashion. Such further specification could, however, be achieved using one degree of freedom planned contrasts (Abelson & Prentice, 1997; Hager, 2002; Niedenthal, Brauer, Robini, & Innes-Ker, 2002). This is especially true for cases where the process is not completely interrupted (i.e., the equivalent of partial tradition mediation).
The interruption of the process, as encoded in $B$, can be achieved in two different ways: either the hypothesized causal path from $X$ to $TV$ can be interrupted (Figure 3(ii), as in Valdesolo & DeSteno, 2008), or the path from $TV$ to $Y$ can be interrupted (Figure 3(iii), as in Zanna & Cooper, 1974). One could also interrupt both paths at the same time, but two distinct possibilities can be considered when planning a test of a process account based on TPIS: Interrupting with the effect of $X$ on $TV$ or with the effect of $TV$ on $Y$. Both tactics must lead to the same result if the process account is viable, namely that the effect of $X$ on $Y$ is eliminated or at least substantially reduced in (II$^\prime$).

EQUIVALENTS OF PARTIAL AND COMPLETE MEDIATION IN TPIS

So far we have discussed the prototypical case where an effect found under $B = -1$ is in fact entirely eliminated in the interrupted process condition $B = +1$. However, as in the distinction between complete mediation and partial mediation in traditional mediation test logic (Shrout & Bolger, 2002; see also Baron & Kenny, 1986), TPIS can also capture a situation where several processes are responsible for the effect of $X$ on $Y$. If a focused process is one of several mediating processes at work (Hayes, 2009; Preacher & Hayes, 2008b), then interrupting it may only attenuate, rather than eliminate, the effect of $X$ on $Y$. This results in an interaction pattern where the effect in the interrupted process condition (II$^\prime$) is not nil (as discussed above for the prototypical case), but merely substantially smaller than in the free process condition (I$^\prime$).

Similarly, the manipulation a researcher implements in order to interrupt the process may not interrupt it completely, but only partially. Then, the difference between the original effect and its remainder with the interruption in place may also be smaller than the original effect (i.e., the latter is not completely, but partially eliminated). As a third possibility, also in line with the assumption that several processes could work simultaneously in producing the effect of $X$ on $Y$, there may be a second process, smaller in its effect than the one in focus, but operating in the reverse direction (Hayes, 2009; Preacher & Hayes, 2008b). This second process would only become apparent once the focal process is interrupted. Such a case would be indicated by a significant interaction in which the effect found in the conditions where the process is not interrupted ($B = -1$) has the opposite sign of the effect in the interrupted process conditions ($B = 1$). This latter pattern would indicate that the two processes of different strength and direction suppress each other, albeit to different degrees (see MacKinnon et al., 2000).

PRECONDITIONS AND CAVEATS FOR TPIS

Spencer et al. (2005) have pointed out that under different circumstances, different approaches may lend themselves to empirical study of processes with different degrees of ease and readiness. In the same spirit, we add that while TPIS may be very well suited to some situations, in other situations traditional mediation analysis or an Experimental-causal-chain design is a better option. A comprehensive and integrated research program will ideally combine different approaches in different studies. Limitations of these various strategies should therefore be taken under consideration in study design choices rather than treated as categorical arguments for or against the use of any particular strategy in general.

If TPIS is to have practical value for a researcher wishing to test a process account, by definition it must of course be feasible to interrupt the hypothesized process by manipulation. For example, in the Valdesolo and DeSteno (2008) example, one must be able to trust that the cognitive load manipulation indeed eliminates differences of elaboration on judgments about other versus self. In this specific example the cognitive load manipulation is sufficiently plausible as well as empirically validated to justify such trust. In other cases, it may be advisable to conduct a pretest in which one confirms that the manipulation to be used in the main study indeed interrupts the process. At first glance such a pretest may seem one must be able to trust that the cognitive load manipulation indeed eliminates differences of elaboration on judgments about other versus self. In this specific example the cognitive load manipulation is sufficiently plausible as well as empirically validated to justify such trust. In other cases, it may be advisable to conduct a pretest in which one confirms that the manipulation to be used in the main study indeed interrupts the process. At first glance such a pretest may seem
process is manipulated. Therefore, a pretest can serve an important function in securing that this manipulation is valid.

The manipulation of $B$ in TPIS must of course be valid by some criterion. Such independent reassurance of validity is a desideratum for virtually any manipulation. In fact, it is also required, in principle, for any measured variable, especially measured mediator candidates. After all, no single study can stand on its own but must instead be viewed in the context of previously obtained knowledge (Kenny, 2008). Whether the required validity results from past research, a pretest designed especially for the focal study that tests a process account, or even a strong plausibility argument (see Sigall & Mills, 1998) that is widely accepted among researchers in a community, is of little importance here. Of course, a particular study may claim to follow TPIS, but actually employ an extremely dubious and shoddily contrived manipulation to interrupt a process. Similarly, a study may draw bold conclusions from a traditional mediation analysis in which the mediator is an unreliable ad hoc measure, has little or no construct validity or even lacks sufficient discriminant validity to be considered theoretically distinct either from $X$ or from $Y$ (Kashy et al., 2009; Spencer et al., 2005). But the possibility of inadequate use does not inherently undermine the validity of any particular design strategy in principle – neither that of TPIS, nor that of traditional mediation analysis (Baron & Kenny, 1986) and its later variants (see Preacher & Hayes, 2008b).

As a terminological note, we add that while it may be tempting to refer to TPIS as “manipulating the mediator,” this phrase is inaccurate.\(^8\) The careful introduction of the manipulated variable $B$ in the formal argument above (see also Figure 3) underlines that it is more accurate to speak of “manipulating whether the process can run freely or not” or “interrupting the process.” The manipulation represented by $B$ does not directly target the transmitting construct (as in Valdesolo & DeSteno, 2008). It encodes different manipulated conditions (i.e., ($\Gamma$) vs. ($\Gamma'$)) rather than a variable at a particular point on a causal path (i.e., $TV$).

Additionally, and related to this issue, TPIS does not simply redefine the terms mediator and moderator to mean the same thing. The variable $B$ introduced above is a different variable than the transmitting construct specified as a traditional mediator. By interfering with the process via TV, $B$ interacts with the manipulation of the original independent variable $X$ in producing an effect, but $B$ does not transmit that effect. Of course, one could include a measurement of $TV$ – similar to a manipulation check for $B$ – in the study design. This would amount to a design of moderated mediation or mediated moderation (Baron & Kenny, 1986; Muller et al., 2005; Preacher, Rucker, & Hayes, 2007). However, it has long been acknowledged that while manipulation checks can be very useful, they also have disadvantages (Sigall & Mills, 1998). Further, our starting point for the present discussion was to avoid measuring the $TV$. Thus, TPIS is especially suited in cases where manipulations, such as cognitive load in the example of Valdesolo and DeSteno (2008), are empirically well established and have high face validity.\(^9\)

### Similarities and Differences of TPIS with and to the MacArthur Approach

The distinction between $B$ and $TV$ makes it clear that it is important to discuss recent work presenting what has become known as the MacArthur approach to mediation and moderation (Kraemer et al., 2002; Kraemer et al., 2008). A detailed presentation of this approach is beyond the present scope, but we discuss three main aspects here. First, the MacArthur approach stipulates conditions central to determine if a third variable is a moderator or a mediator. In particular, according to the authors, a moderator candidate must be unrelated to an independent variable and “precede [it]” (p. S103) to be eligible as a moderator at all. The MacArthur approach and traditional moderation analysis thus share a basic conceptual understanding of a moderator, but the MacArthur approach differs in the criteria imposed on study designs in order to secure unambiguous conclusions regarding the moderating role of a third variable. The logic behind TPIS is also based on an orthogonal manipulation (and thus independence) of $X$ and the factor $B$. However, the variable $B$ in TPIS is not a theoretical variable in itself and should not be considered as a moderator of theoretical and psychological import in the sense of Baron and Kenny (1986). Rather, $B$ is a technical factor arising from the logic of the experimental design that captures whether the hypothesized process is uninterrupted or interrupted by manipulation. Thus, the MacArthur approach aims to modify and sharpen the criteria for a variable to be viewed as a moderator whereas TPIS uses an orthogonal, interacting factor $B$ as a tool in the service of examining process hypotheses. This factor $B$ does not possess psychological substance as does a moderator in the traditional Baron and Kenny (1986) sense.

Second, Kraemer et al. (2002, 2008) extend the concept of mediation as introduced by Baron and Kenny (1986). In particular, they propose that “[t]reatment [i.e., $X$; J&S] may not merely change the level of $M$ […] but may change the nature of $M$” (Kraemer et al., 2002, p. 879). In our reading, this suggests that the concept assessed as a $TV$ differs between conditions, as a result of the manipulation of $X$. Also, the MacArthur approach views a variable as a mediator whose variance or relationship to the dependent variable vary as a function of the independent variable (Kraemer et al., 2008). TPIS in its present scope, in contrast, somewhat conservatively affirms the traditional concept of a process inherent in mediation analysis where different levels in $X$ predict different levels in $TV$. These different levels in $TV$, in turn, predict

\(^8\)We also propose that the common expression “measuring a process” as a synonym for the traditional mediation approach is similarly misleading. What is measured in a mediation following Baron and Kenny (1986) is not the process, but the state of a variable in the course of the process. Given appropriate and plausible assumptions (Sobel, 2008), this measurement allows the inference that there has been a causal process at work. However, the measurement and a significant indirect effect do not constitute the observation of the process proper (see also Sigall & Mills, 1998 and Spencer et al. 2005, for a similar discussion of the relation between psychological constructs and observables).

\(^9\)One could argue that, strictly speaking, a test of validity for the manipulation of $B$ in each instance of application of TPIS is necessary. While we do not disagree, we note that parallel requirements could be imposed on traditional mediation analysis. Specifically, one could strictly insist upon an independent reassurance of the construct validity and reliability of a measured $TV$ in every instance of traditional mediation analysis (Baron & Kenny, 1986; Kashy et al., 2009; Kenny et al., 1998; Mathieu & Taylor, 2006).
different levels in Y without a change of the nature, variance, or relationships of a variable along the process.

Finally, the MacArthur approach is primarily concerned with observed and measured third variables as mediator or moderator candidates. It “seek[s] to generate strong hypotheses for future studies” (Kraemer et al., 2008, p. S107), the analyses conducted within the MacArthur approach are “hypothesis-generating rather than hypothesis-testing” (Kraemer et al., 2002). TPIS, in contrast, is a technique for hypothesis testing. In TPIS a researcher intends to empirically scrutinize a process account in a design with a maximal degree of experimental manipulation. In sum, while it has similarities with the MacArthur approach, TPIS represents a distinct strategy to test process hypotheses and targets different questions and concepts than the MacArthur approach.

CONCLUSION

We have shown how a theoretical process account can be translated into a 2 × 2 design and tested via an interaction term. Specifically, following common causal logic, an effect should occur if the process explaining it can run freely, but not if the process is interrupted. TPIS is rooted in this simple idea. If measuring a mediator is problematic or undesirable for some reason, if one wants to experimentally test whether an effect factually vanishes if the suspected process is interrupted by manipulation or if a researcher simply wants to muddle all available design options before deciding on any particular one, TPIS offers two ways to interrupt the effect to be explained. Either the effect of an independent variable on a hypothesized transmitting variable or the effect of this transmitting variable on a dependent variable can be interrupted experimentally.

The examples from the literature mentioned above (as well as those explored in Spencer et al., 2005, or Sigall & Mills, 1998) show that TPIS is actually already used. However, this strategy has not, to our knowledge, been outlined in a detailed way. We have therefore provided a formal account of how a process hypothesis can be cast as a traditional mediation analysis problem as well as an interaction hypothesis with the goal to contribute toward the recognition that processes can be empirically investigated in ways other than traditional mediation analysis. We hope that the general formulation provided here renders TPIS more visible and applicable as a design option so that it can contribute to the advancement of process research.

ACKNOWLEDGEMENTS

We have greatly profited from discussion with and comments on earlier versions of this manuscript by Markus Brauer, Maya Machunksy, Dominique Muller, Benjamin Nagengast, Devin Ray, Thomas Schubert, and Vincent Yzerbyt. Thank you.

REFERENCES


APPENDIX: ORIGINAL TEXT IN THE SURVEY AMONG MEMBERS OF THE SPSP-DISCUSS MAILING LIST

It is an established phenomenon that an immoral behavior is judged as more severe if shown by another person than by oneself. This difference is called the moral hypocrisy effect. The authors of the study in focus here either (a) induced participants to commit a transgression against a fictitious other participant or (b) told participants that another participant had committed the same transgression. Orthogonally, the authors manipulated whether participants made the judgment of the transgression while rehearsing a string of seven digits or they made this judgment without such an extra task. The dependent variable was the degree to which participants rated the behavior (their own or that of the other target’s) as fair. Ninety-one individuals participated in the study and were randomly assigned to one of the resulting four cells of a $2 \times 2$ factorial design. The results of the study are visualized in Figure A1.

The difference between the self and the other judgments within the no load condition was significant, while the difference between the target conditions within the load conditions was virtually zero. This resulted in a significant interaction effect.

Also, for the self, cognitive load led to lower fairness judgments compared to no load, but judgments for the “other” target did not differ as a function of the cognitive load manipulation.

Now, we are interested in which conclusion you believe one can draw from this study.

![Figure A1. Schematic visualization of the interaction effect found by Valdesolo and DeSteno (2008)](image)